

あきたこまち栽培記録に基づくイネの出穂日予測

井内康裕^{†1} 深澤佑介^{†2} 上原宏^{†1} 金田吉弘^{†1}

概要：秋田県内広域観測地点での過去 25 年間のイネの生育記録、および該当生育地点での気象情報をもとに機械学習(XGboost)による出穂日(開花日)の予測を試みた。生育記録を特徴量とした場合、±2.72 日の予測精度が得られた。生育記録は、農業生産者の経験を反映したデータを含んでいる。経験を反映したデータを除外し、気象データの特徴量を含めたところ、予測精度は±2.19 日となり、特徴量としては気温の変動幅の貢献度合いが大きいことが判明した。気温の変動と出穂日との関係を分析したところ、盆地特有な気温変動が出穂日に影響する等、秋田県内の地域特性と出穂日との関係を示唆する結果が得られた。

キーワード：イネ、あきたこまち、出穂日予測、機械学習、XGboost

1. はじめに

本研究では、'あきたこまち'の栽培記録にもとづき出穂日(イネの開花日)の予測および予測を説明づける特徴量のうちで特に地域特性に関連する特徴量に関する分析を試みる。栽培記録および気象データにもとづくイネの出穂日予測は、[1]をはじめとしてすでに様々な研究が行われ、品種別に高い予測精度が得られている。本研究では、秋田県内各地域での過去 25 年間の'あきたこまち'の栽培記録をもとにして、秋田県内各地域の気象特性と出穂日との関連性に着目した分析を試みる。

2. データセット

2.1 あきたこまち生育記録

分析には秋田県内において広域に観測された'あきたこまち'の生育記録[a]を使用した。このデータは、1993 年-2018 年の期間に観測されたフィールドデータであり、サンプル数は 1608 である。

データ項目は、観測住所、観測年、土壌型、標高、移植日(田植え日)、出穂日(開花日)、栽植密度、および生育状態に関する定期観測値として、草丈、莖数、葉数、SPAD(葉の中に含まれる葉緑素量)等が含まれる。定期観測日は全てのデータについて 6 月 10 日、6 月 25 日、7 月 5 日、7 月 15 日である。

データ項目のうち、移植日から出穂日までの日数(以降、出穂日数)を教師データ、それ以外の項目を特徴量とした。

2.2 微気象データ

上記、生育記録の観測住所付近の気象データを気象庁ホームページ[b]から取得した。具体的には、各データの観測年における日次平均気温、および日射時間を取得し、特徴量とした。

3. 分析の考え方

機械学習器 XGboost を用いて、出穂日数の学習・予測を行い、予測精度に対する各特徴量の貢献度を分析する。出穂日は作期や栽培地域に関わる重要な形質[c]とされ、農業生産者はそれぞれの地域における長年の経験をもとにして、移植日を決定する。このように移植日には農家の経験が折り込まれており、これを特徴量に取り込むことで、経験が出穂期間予測にどの程度貢献するかを評価できる。本研究では、経験を反映した特徴量を含めた場合の予測と、経験に依存しない特徴量による予測とを比較し、経験に代替する特徴量が何かについて分析を行った。

イネをはじめとする作物の生育は、気象条件との関連性が大きいことから[1,2,3]、移植日は主に当該地域の気象特性に関する経験から導かれたものと考え、地域別気象特徴量が移植日特徴量にどのような代替性を有するかを中心に分析を行った。

4. 分析結果

4.1 移植日の予測に対する貢献度

生育記録のうち、出穂日を除く全データ項目を特徴量として、出穂期間の予測を行った。予測精度(RMSE)は、±2.72 日であった。図 1 に予測貢献度の高い特徴量を示す。移植日の貢献度が突出して高いことがわかる。

4.2 移植日に代替する特徴量の探索

上記の特徴量から移植日を除外した場合、予測精度は ±4.05 となり、予測貢献度が最大の特徴量は標高であった。次に、気象に関する特徴量を加味することで移植日を特徴量に加味した場合と同等な予測精度を確保できるかについて分析を行った。

まず移植日から 60 日、70 日、80 日それぞれにおける日次平均気温の積算値および 1 日あたり日射時間積算値を特徴量

^{†1} 秋田県立大学
Akita Prefectural University

^{†2} 東京大学 人工工学研究センター
Tokyo University Research into Artifacts, Center for Engineering

a) 秋田農林水産部提供による

b) <https://www.data.jma.go.jp/obd/stats/etrn/index.php>

c) 国立研究開発法人 農業生物資源研究所
<http://www.naro.affrc.go.jp/archive/nias/org/DivPlant/Functional/page01.html>

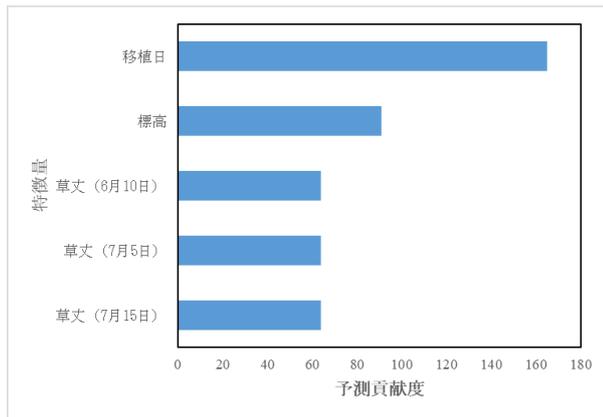


図1 栽培記録特徴量の予測貢献度

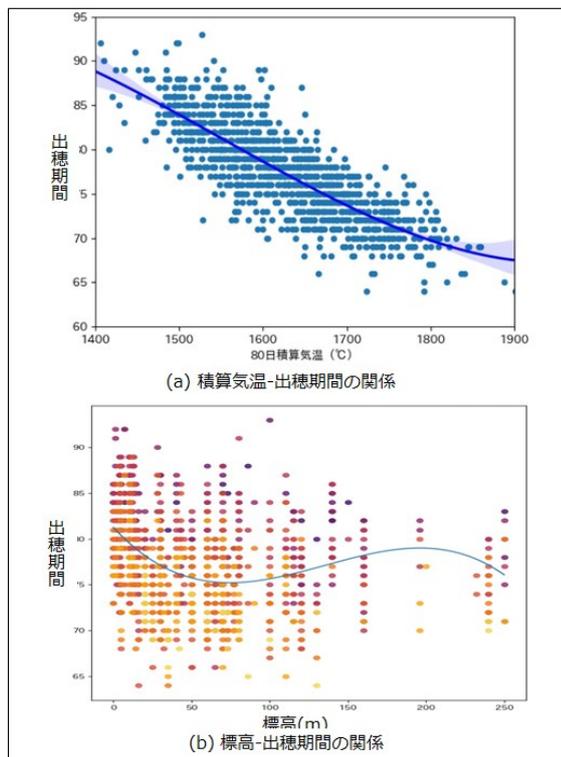


図2 積算気温、標高と出穂期間との関係

に加えたところ、予測精度は、 ± 2.72 となり、予測貢献度が最大の特徴量は標高、次いで積算気温（80日積算気温）であった。一方、日射時間積算値は相対的に貢献度が小さい値となった。図2に出穂日に対する積算気温、標高それぞれの回帰曲線を示す。

図2(a)のとおり、積算気温と出穂日はほぼ線形となる。標高と気温は単調減少の関係にあるため、標高が積算気温を代替する特徴量であるならば、標高と出穂日も単調な関係となる。しかし、図2(b)が示す通り、両者は単調ではない。図2(b)のデータ点の配置を地域別に分析したところ、標高50m付近で出穂期間が顕著に小さいデータは盆地（横手市、大仙市）が支配的であることがわかった。盆地は一般に季節的な寒暖差が大きいことから、標高は、積算気温

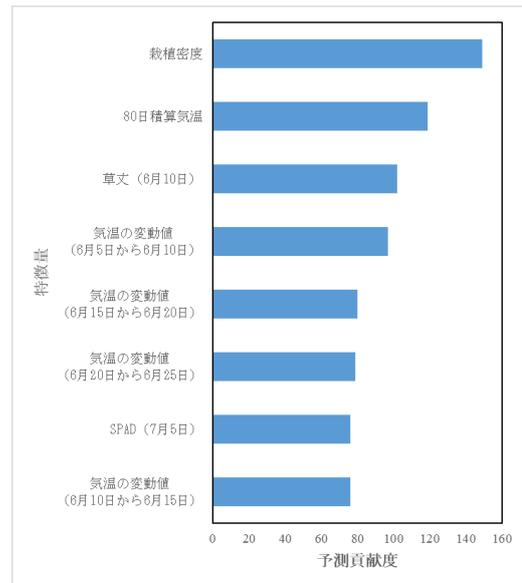


図3 気象データを加味した予測貢献度

および気温変化量の出穂期間に対する相関が混在した特徴量となっていることが想定される。

そこで、積算気温に加えて、5日毎の気温差を特徴量としたところ、予測精度は ± 2.21 に向上した。更にこの特徴量から標高を除外すると、予測精度は ± 2.19 となり、気温差を加味すると、標高が予測精度に対して寄与しなくなることを示唆する結果が得られた。図3に主な特徴量を示す。

6月における5日毎気温差が貢献度の上位を占めていることがわかる。

5. おわりに

移植日は、出穂日に関する各地域の経験によって決定されるもので、XGboostによる学習・予測からは、その経験を支持する結果を得た。移植日を除外し、積算気温および気温差を特徴量に加えることで、移植日を含む場合を上回る予測精度が得られた。図3によると、特定の期間での気温差が予測精度に貢献しており、気温差の系列性を示唆していることから、今後、気温差の系列性、すなわち地域毎の気温変動パターンを考慮した特徴量を用いることで、出穂日と地域特性との関連性を明確化したい。

参考文献

- [1]堀江武，中川博視，‘イネの発育過程のモデル化と予測に関する研究-1-モデルの基本構造とパラメータの推定法および出穂予測への適用’日本作物学会紀事 59(4), pp687-695(1990).
- [2]丹野耕一，高橋重郎，千葉隆久，高橋正道，‘水稻における出穂期の推定について—移植期と出穂期の関係—’，東北農業研究, 23, pp. 17-18(1978).
- [3]近藤拓也，西内俊策，‘イネの出穂予測精度向上を目指した農業情報の利用’，第31回人工知能学会全国大会(2017).